

AUTOMATIC TUNING OF A GRAPH-BASED IMAGE SEGMENTATION METHOD FOR DIGITAL MAMMOGRAPHY APPLICATIONS

Hirota Susukida, Fei Ma *, Mariusz Bajger

School of Informatics and Engineering, Flinders University
P.O. Box 2100, Adelaide SA 5001, Australia

ABSTRACT

Mammogram segmentation tasks underpin a wide range of registration, temporal analysis and detection algorithms. Unfortunately, finding an accurate, robust and efficient segmentation still remains a challenging problem in mammography. A recent segmentation technique, based on minimum spanning trees (MST segmentation), is known to be robust to typical mammogram distortions and computationally efficient. This method captures both local and global image information but the balance requires choosing a parameter. So far no automatic procedure to estimate this parameter has been proposed and the value was determined experimentally. In this paper a segmentation evaluation criterion, based on a measure of image entropy, is used to automatically optimize the granularity of an MST-based segmentation. The method is tested on a set of 82 random images taken from a commonly used mammogram database. The results show a dramatic improvement in the accuracy of a MST segmentation tuned up using the entropy-based criterion.

Index Terms— Image segmentation, entropy, minimum spanning tree, mammography

1. INTRODUCTION

Diagnostic imaging technologies such as magnetic resonance imaging, computed tomography and digital mammography provide a vast set of valuable tools in modern medicine. These tools are often critical in both diagnosis and treatment planning. In particular, image segmentation algorithms allow for computer-automated delineation of regions of interests - often, but not always, anatomical structures - and play a vital role in a large number of biomedical imaging applications [1].

Despite numerous techniques that have been proposed in pursuit of an adequate segmentation method in the field of digital mammography there is still no exact solution to this complex problem. The complexity of mammograms comes from inherent blurring caused by round anatomical feature shapes in the direction of X-ray beam and superimposed boundaries resulting from overlapping features in the path of each X-ray beam. Thus to determine an accurate boundary of an image feature it is necessary to utilize both local and global image information in the segmentation algorithm. If we also add a requirement of computational efficiency, and take into account the large size of a mammogram, the difficulty of this task is evident.

A promising segmentation algorithm which uses both local and global image information, and is sufficiently efficient for practical applications, has been proposed in [2]. The algorithm considers an image as a graph and builds minimum spanning trees to segment the

image components (MST segmentation). The technique is particularly interesting in the light of a recent research showing its applicability to mammogram segmentation [3], mammogram registration and temporal analysis of mammograms [4].

The coarseness of MST segmentation is controlled by a single parameter k , which has to be determined a priori. Finding the best value of k for a large class of images is a challenging task and no automatic method has been proposed so far. Using an entropy-based image evaluation method developed in [5], we show that MST segmentation can be optimized for large classes of mammograms. We test our findings on a sample of 82 mammograms taken from a commonly used publicly available mammographic database [6]. Superiority of the segmentation with k estimated using an image entropy measure is evident. The improvement is rigorously measured in context of pectoral muscle line detection (see e.g. [7], [3]). The results are very promising: for 84% of tested images a significant improvement (in comparison to experimentally chosen value of k) of segmentation accuracy, at least 20%, was observed.

Pectoral muscle appears as a bright triangular patch in the upper left corner of the mammogram (see Figures 1, 3). It is an important feature of mammograms often used as a landmark for image registration (see e.g. [7], [3]).

2. MST-BASED IMAGE SEGMENTATION METHOD

For detailed description of the algorithm in context of mammography we refer the interested reader to [3], and for general discussion to [2]. For the convenience of the reader we briefly describe the vital steps.

Let $G = (V, E)$ be an undirected graph such that V (the set of vertices) is the set of pixels in the image and E is the set of edges that connect pixels to immediate neighbors. Each edge is assigned a weight defined as follows:

$$w((v_i, v_j)) = \begin{cases} |I(v_i) - I(v_j)|, & (v_i, v_j) \in E, \\ \infty, & \text{otherwise,} \end{cases}$$

where $I(v_i)$ is the image intensity at v_i . A tree that spans a component $C \in V$ and has a minimum total weight is called a *minimum spanning tree* of C (MST(C)). A segmentation is a partition of the set V in a graph $G' = (V, E')$, where $E' \subseteq E$. For a component $C \in V$, the *internal difference*, $Int(C)$ is defined as the largest weight in the MST(C). For two components C_1 and C_2 , their difference $Diff(C_1, C_2)$ is defined as the minimum weight edge connecting C_1 and C_2 .

The segmentation process starts from the trivial partition with each pixel (vertex) being a single component. The components C_1 and C_2 are merged if

$$Diff(C_1, C_2) \leq \min(Int(C_1) + \tau(C_1), Int(C_2) + \tau(C_2)).$$

*Thanks to National Breast Cancer Foundation for funding.

The threshold function τ is given by $\tau(C) = \frac{k}{|C|}$, where $|C|$ stands for the number of elements in component C , and k is a constant.

The constant k is the only runtime parameter used in MST segmentation algorithm. It controls the degree of similarity between the components and hence the number and size of components found.

3. ENTROPY-BASED SEGMENTATION EVALUATION METHOD

The following function of an image I , has been proposed in [5] as a measure of effectiveness of an image segmentation:

$$E(I) = H_l(I) + H_r(I). \quad (1)$$

The first term - the *layout entropy* - measures the global image disorder (generally it increases with the number of components), and is defined by the formula

$$H_l(I) = - \sum_{j=1}^N \frac{S_j}{S_I} \log \frac{S_j}{S_I},$$

where S_I is the area of the whole image and S_j is the area of the j -th component.

The second term - the *region entropy* - measures the uniformity within components (it decreases when the number of regions increases), and is given by the formula

$$H_r(I) = \sum_{j=1}^N \frac{S_j}{S_I} H_v(R_j),$$

where $H_v(R_j)$ is the entropy for region j for the pixel property (a feature) v . In [5], luminance was used as the feature v . In this paper we use pixel's intensity value as v . Denoting by $V_j^{(v)}$ the set of values associated with feature v in region j and by $L_j(m)$ the number of pixels in region j with value m for feature v , the j -th region entropy is expressed as

$$H_v(R_j) = - \sum_{m \in V_j^{(v)}} \frac{L_j(m)}{S_j} \log \frac{L_j(m)}{S_j}.$$

In what follows we will show that function E can be effectively used to tune up MST segmentation of mammograms.

4. EXPERIMENTS

The MST segmentation was tested in [2] on natural scenery images and images from Columbia COIL database. Medical images - mammograms - were successfully segmented with the MST method in [3], in pursuit of an exact boundary of the pectoral muscle. The major difficulty in application of the method is the choice of a single parameter value k which controls the granularity of the segmentation. Experimentally chosen values, although worked fairly well in [2] and [3], were certainly not optimal for many images. The differences can be very significant (Figure 1).

As pointed out in [5], the difficulty of finding an optimal segmentation comes from the lack of tools allowing for an objective judgment of a segmentation goodness for a large class of images. Needless to say, the problem is particularly hard for medical images such as mammograms due to a very limited perceptual information available to the algorithm or a human marker. One of the rare situations in mammography when the boundary of a region of interest

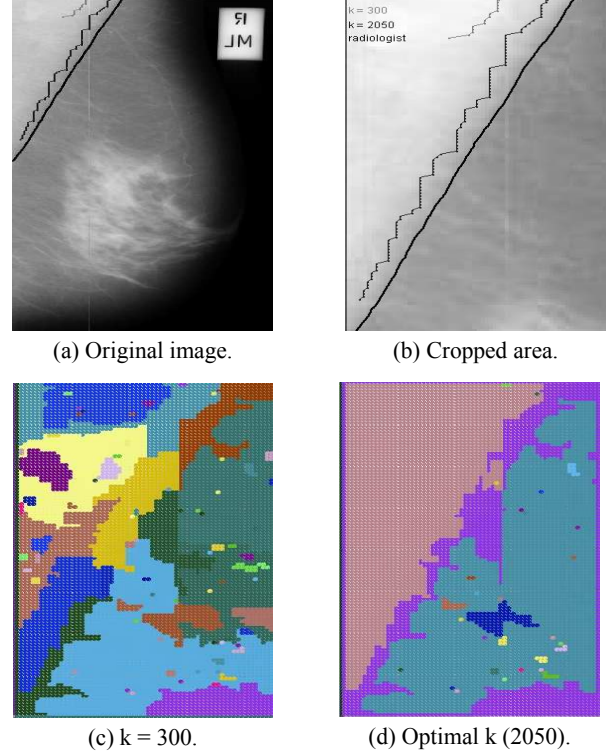


Fig. 1. Pectoral muscle lines and segmentation results for two different values of k : $k = 300$ and $k = 2050$ (image mdb042). (a), (b) The thick and smooth black line is the line drawn by a radiologist, and the other two thin ragged lines are found by an MST segmentation. The closer to the radiologist line, and longer, is the line found for the optimal k value (see Section 4). The most left lighter line is the one found for k set to 300. (c), (d) The difference in segmentation granularity for the two values of k .

is fairly well determined (albeit not always) is the pectoral muscle area. For this reason we choose pectoral muscle detection task to estimate usefulness of entropy-based tuning of MST segmentation. However, our ultimate goal is to be able to produce a reliable segmentation of the whole mammogram. As shown in [4], MST segmentation can provide robust segmentation, that is, stable with respect to small shifts and rotations of the image. In what follows we show that MST segmentation of the pectoral muscle, can be very accurate, thus making the process useful for mammogram registration and mammogram temporal analysis tasks.

4.1. Data

In order to determine the performance of the proposed schema 82 images were selected from the Mini-MIAS database of mammographic images [6]. The same set was used in [3] and [7]. The spatial resolution of these images is $200 \mu\text{m}$ and a depth resolution is 8 bit. The images in the database are 1024×1024 pixels in size. To reduce the computation time, before applying the MST algorithm, the images were automatically cropped to a rectangle enclosing the pectoral muscle region, and subsampled by a factor of four.

The coordinates of the radiologist lines used in this study are the same as used in [3] and [7]. The lines were drawn by an experienced

radiologist and kindly provided to the authors by R. J. Ferrari and R. M. Rangayyan.

4.2. The algorithm

Each image was segmented with the MST method and the image entropy was calculated for each k value within the range $[300, end]$, with step = 50 (that is, $k = 300, 350, 400$, and so on). The range for k was determined as follows. Number 300 was chosen as the starting value for comparison with results given in [3] where it was experimentally determined that 300 provides reasonable results to all 82 images. The upper range boundary end was set to $\frac{1}{3}$ of the total number of pixels in the cropped image. This value was determined based on the considerations in [2] concluding that the value of k effectively estimates an average size of the segmented components. That is, only salient components smaller than a given value of k can survive the merging process and become segmented as stand-alone components. Since we are looking for the component enclosing the pectoral muscle area which for each correctly cropped image sits in the upper left corner (right breast images are flipped to left), occupies approximately half of the cropped image area, and is one of the most salient large regions in the image, one third of the image size seems a reasonable choice (see 5.2 for more discussion).

The experiments were conducted twice: once using raw images and again on images smoothed with a one-dimensional Gaussian filter with $\sigma = 0.5$ and oriented at an angle $\theta = \frac{\pi}{4}$ to approximate the expected angle of the pectoral muscle boundary. It is known ([2], [3]) that MST segmentation is generally more accurate when an image is smoothed and the same smoothing was used in [3].

4.3. Results

To judge the accuracy of the segmentation we focused on detection of pectoral muscle boundary. We measured the improvement in terms of distance between a gold standard - the radiologist drawn line - and the lines obtained for two values of k : $k = 300$, used in [3] for all 82 images, and the k value determined by the image entropy measure E defined in (1). The distance is calculated as the area (number of image pixels) between each of the obtained lines and the gold standard. The following formula is used to measure the improvement

$$improvement = \frac{A_{300} - A_{opt}}{A_{300}} * 100\%, \quad (2)$$

where A_{300} denotes the area between the radiologist line and the line found for $k = 300$, and A_{opt} denotes the area between the radiologist line and the line found for k determined with the entropy measure.

Note that the maximum improvement - 100% - happens when the new line entirely covers the radiologist line ($A_{opt} = 0$), and no improvement if $A_{opt} = A_{300}$. Possible deterioration is indicated by a negative value. To calculate the areas the biggest overlapping range of values is taken for all tree lines. (This may result in some underestimated improvement, see 5.1 for more discussion.)

Table 1 summarises the results. The first row shows the statistics for raw images and the second for Gaussian smoothed images. In 7 cases the improvement was negative (first column) which means that the line found for $k = 300$ was closer to the radiologist line. In 8 cases (10 for Gaussian smoothing) the results were the same. In 3 cases a small (up to 20%) improvement was noted and in 64 cases (78% of all images) at least 20% improvement was recorded. The average improvement was 40% for raw images and, slightly less,

Table 1. Entropy tuned MST segmentation improvement.

$filter \backslash \%$	-	0%	0-20%	20-50%	50%+	Mean
no	7	8	3	27	37	40%
yes	7	10	3	24	38	36%

36% for smoothed images. This confirms the fact that MST segmentation generally works better for smoothed images, hence the smaller overall improvement value for those images. It is also worth mentioning that in almost half of the cases the improvement was 50% or more, that is, the new line was at least two times closer to the radiologist line.

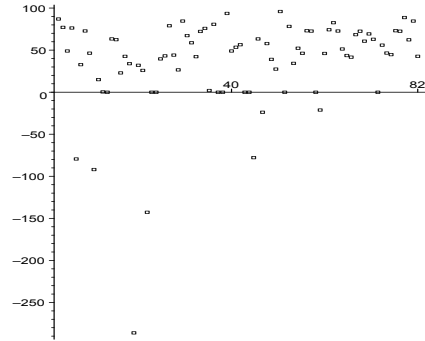


Fig. 2. Improvement values for each of the 82 images.

5. DISCUSSION

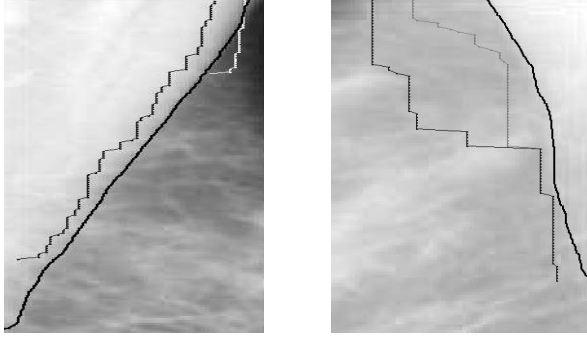
5.1. Negative improvement cases

There were 7 cases for which the entropy method, did not help with the pectoral line detection (see Figure 2). In fact, the error was as big as 286% for mdb046 image. These cases can be classified as false failures (mdb046, mdb037, mdb107, mdb091) and true failures (mdb033, mdb049, mdb089). Figure 3 shows examples of both kinds. False failures are cases where the formula produces a negative value but the line found for an optimal value of k is in fact better than the one found for $k = 300$. Unfortunately, the evident superiority of the longer line (Figure 3 (a)), is not reflected in (2) since only the overlapping range for both lines counts in (2).

True failures, like mdb049, were caused by weak pectoral muscle boundary line, presence of a competitive bright region in the image (for very dense breasts) or very low saliency of the pectoral muscle area (e.g. mdb033). Summarising, in three cases (out of 82) the entropy-based MST segmentation tuning did not work well.

5.2. k value range

In 54 cases the entropy of the image attained its global minimum for a value of k inside the range (typically several local minimum were present) (see Figure 4). In 22 cases the minimum was attained inside and stayed on the same level for the rest of the range, and in the remaining 6 cases the global minimum happened at the end of the range, which suggest possible improvement for these images. We further analysed these 6 images.



(a) Image mdb046 - false failure. (b) Image mdb049 - true failure.

Fig. 3. Difficult cases. (a) Despite the negative improvement value resulting from (2), the line for optimal value of k is longer and better approximates the (thick black) radiologist line than the one found for $k = 300$ (very short lighter line, first from the right). (b) The algorithm failed. The ragged long line for optimal k is far from the radiologist line.

All but two (mdb033 and mdb091) of them have shown very significant improvement (70% or more). The mdb091 case belongs to false failures category, as discussed in 5.1, while mdb033 is a true failure. It is a particularly difficult image to process due to a very dense breast - a strong competitive component present - and a very low saliency pectoral muscle area. Range increments, up to $\frac{1}{2}$ of the image size, have not improved the line detection process.

Figure 4 shows an example of a most typical entropy behaviour for different k values (image mdb076). The improvement, measured by formula in Eq. (2), for this image was 76%.

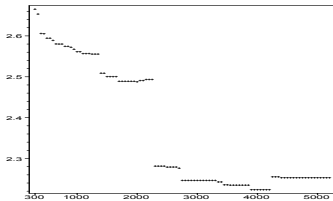


Fig. 4. Entropy vs k value for mdb076.

5.3. Comparison with other published results

Since the same data were used in [3] it is tempting to compare our findings with those presented in the paper. There are however major differences between the two approaches. The performance of MST segmentation was evaluated in [3] by measuring the area between the radiologist drawn boundary and the predicted boundary in terms of numbers of false positives (FP) and false negatives (FN). FP were the pixels assigned by the algorithm to the pectoral muscle area but assigned by the radiologist to outside the pectoral muscle, and FN were defined as the pixels assigned by the algorithm to the outside of the pectoral muscle area but assigned by the radiologist to the pectoral muscle area. This approach is hardly adequate for the entropy-tuned MST segmentation evaluation since if the radiologist line is shorter than the line predicted by the algorithm - which happened on several occasions (see Figures 1, 3) - the extra line points present were counted as false positives, thus penalizing the additional boundary

Table 2. Comparison with [3]. First two columns show the average proportion of FP and FN scores for 82 images.

	FP	FN	FP and FN less than 0.05
in [3]	0.0255	0.1168	40 images
here	0.0561	0.0821	44 images

found by the algorithm. This is a reasonable approach for the pectoral muscle line detection but is certainly questionable for other applications where invisible to a human boundary may be well determined (in terms of usability for applications) by an algorithm. This includes mammogram analysis since radiologists most often cannot determine mass boundaries at all. Such contours are however necessary for discrimination between benign and malignant masses ([8]), image registration tasks and many other biomedical-imaging applications ([1]).

Table 2 compares the results from [3] with the one obtained here using entropy-tuned MST segmentation. The bias towards FP is evident. We have to mention that the results in Table 2 concern the lines found by an active contour model (snake) algorithm implementation ([9]), with MST segmented regions used as initial snake contours.

6. REFERENCES

- [1] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annu. Rev. Biomed. Eng.*, vol. 2, pp. 315–337, 2000.
- [2] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vision*, vol. 59, no. 2, pp. 167–181, 2004.
- [3] F. Ma, M. Bajger, J. P. Slavotinek, and M. J. Bottema, "Two graph theory based image segmentation methods for identifying pectoral muscle in mammograms," *Pattern Recognition*, vol. 40, pp. 2592–2602, 2007.
- [4] F. Ma, M. Bajger, and M. J. Bottema, "Robustness of two methods for segmenting salient features in screening mammograms," in *Proc. of the 9th Conf. on Digital Image Computing Techniques and Applications (DICTA2007)*, Adelaide, 2007.
- [5] H. Zhang, J. E. Fritts, and S. A. Goldman, "An entropy-based objective evaluation method for image segmentation," in *Proc. of SPIE: Storage and Retrieval Methods and Applications for Multimedia*, 2004, vol. 5307, pp. 38–49.
- [6] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Bental, and J. Savage, "The mammographic images analysis society digital mammogram database," *Excerpta Medica*, vol. 1069, pp. 375–378, 1994.
- [7] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, R. A. Borges, and A. F. Frère, "Automatic identification of the pectoral muscle in mammograms," *IEEE Trans. Med. Imaging*, vol. 23, no. 2, pp. 232–245, 2004.
- [8] C. Varela, S. Timp, and N. Karssemeijer, "Use of border information in the classification of mammographic masses," *Phys. Med. Biol.*, vol. 51, pp. 425–441, 2006.
- [9] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *Int. J. Comput. Vision*, vol. 1, no. 4, pp. 321–331, 1987.